Appl Microbiol Biotechnol (2013) 97:4289–4300
DOI 10.1007/s00253-013-4865-2

MINI-REVIEW

Genome-wide analysis of the $Zn(II)_2Cys_6$ zinc cluster-encoding gene family in *Aspergillus flavus*

Perng-Kuang Chang · Kenneth C. Ehrlich

Received: 24 January 2013 / Revised: 13 March 2013 / Accepted: 18 March 2013 / Published online: 6 April 2013

Abstract Proteins with a $Zn(II)_2Cys_6$ domain, Cys-X_2-Cys-X_6-Cys-$X_{5\text{-}12}$-Cys-X_2-Cys-$X_{6\text{-}9}$-Cys (hereafter, referred to as the C6 domain), form a subclass of zinc finger proteins found exclusively in fungi and yeast. Genome sequence databases of *Saccharomyces cerevisiae* and *Candida albicans* have provided an overview of this family of genes. Annotation of this gene family in most fungal genomes is still far from perfect and refined bioinformatic algorithms are urgently needed. *Aspergillus flavus* is a saprophytic soil fungus that can produce the carcinogenic aflatoxin. It is the second leading causative agent of invasive aspergillosis. The 37-Mb genome of *A. flavus* is predicted to encode 12,000 proteins. Two and a half percent of the total proteins are estimated to contain the C6 domain, more than twofold greater than those estimated for yeast, which is about 1 %. The variability in the spacing between cysteines, C_3-C_4 and C_5-C_6, in the zinc cluster enables classification of the domains into distinct subgroups, which are also well conserved in *Aspergillus nidulans*. Sixty-six percent (202/306) of the *A. flavus* C6 proteins contain a specific transcription factor domain, and 7 % contain a domain of unknown function, DUF3468. Two *A. nidulans* C6 proteins containing the DUF3468 are involved in asexual conidiation and another two in sexual differentiation. In the anamorphic *A. flavus*, a homolog of the latter lacks the C6 domain. *A. flavus* being heterothallic and reproducing mainly through conidiation appears to have lost some components involved in homothallic sexual development. Of the 55 predicted gene clusters thought to be involved in production of secondary metabolites, only about half have a C6-encoding gene in or near the gene clusters. The features revealed by the *A. flavus* C6 proteins likely are common for other ascomycete fungi.

Keywords *Aspergillus flavus* · Zinc-cluster protein · Genome · Gene cluster · Secondary metabolite · DUF3468

Introduction

Biological systems contain various groups of DNA-binding proteins that are involved in regulation of many vital cellular processes, such as DNA replication, DNA repair, recombination, and transcription control. The most commonly known DNA-binding proteins include those termed zinc finger, helix-turn-helix, helix-loop-helix, basic leucine zipper, and high mobility group box, which are characterized by the secondary structure of their DNA-binding motifs. The zinc-binding proteins form one of the largest families of transcription factors in eukaryotes. In general, they are categorized into three main classes based on their zinc finger binding motifs (MacPherson et al. 2006), i.e., Cys_2His_2 (C2H2), Cys_4 (C4), and Cys_6 (C6). Only fungi and yeast contain the C6 zinc cluster DNA-binding proteins; this class of proteins hasn't been found in bacteria, plants, and animals. This review summarizes roles of known fungal C6 proteins and deciphers features of C6-encoding genes in the *Aspergillus flavus* genome including subgroups of the C6 domains, functions of a unique domain, DUF3468, and physical association of C6 domain genes with the predicted 55 secondary metabolite gene clusters.

Gal4p, the classical model of C6 zinc cluster DNA-binding protein

The best studied C6 protein is the Gal4 transcriptional activator of the budding yeast, *Saccharomyces cerevisiae*

Electronic supplementary material The online version of this article (doi:10.1007/s00253-013-4865-2) contains supplementary material, which is available to authorized users.

P.-K. Chang (✉) · K. C. Ehrlich
Southern Regional Research Center, Agricultural Research Service, U. S. Department of Agriculture,
1100 Robert E. Lee Boulevard,
New Orleans, LA 70124, USA
e-mail: perngkuang.chang@ars.usda.gov

(Johnston 1987). Gal4p binds to four related 17-base-pair sequences within an upstream activating sequence to activate transcription of the Gal1 and Gal10 genes that are required for catabolism of galactose. Studies have identified various functional domains in the 881-amino-acid Gal4 protein; they include a DNA-binding domain (residues 1–65) (Keegan et al. 1986), a dimerization domain (residues 65–94) (Hidalgo et al. 2001; Hong et al. 2008), and three acidic activation domains (Ma and Ptashne 1987b), and a region near the C-terminus that binds the inhibitor Gal80p (Ma and Ptashne 1987a). The six cysteine residues bind to two Zn(II) ions in a bimetal-thiolate cluster (Pan and Coleman 1990), and the term "binuclear-cluster zinc-finger" DNA-binding domain is used interchangeably. Commonly, $Zn(II)_2Cys_6$ DNA binding domains interact with DNA binding sites consisting of conserved terminal trinucleotides, which are usually in a symmetrical configuration and are spaced by an internal variable sequence of defined length ranging from 2 to 17 nucleotides (MacPherson et al. 2006; Todd and Andrianopoulos 1997).

Functions of previously characterized fungal C6 proteins

The earliest studied fungal C6-type zinc cluster proteins belonged almost exclusively to the ascomycete family (*Ascomycota*) of fungi, such as *Aspergillus nidulans* and *Neurospora crassa* (Todd and Andrianopoulos 1997). Only one has been reported from *Basidiomycota* (Endo et al. 1994) and none from *Chytridiomycota* and *Mucoromycotina*. Characterized fungal C6 proteins have the basic structure of yeast Gal4p except for the Gal80p-binding acidic region (Fig. 1). The publicly available fungal genome sequences at the Broad Institute (http://www.broadinstitute.org/annotation/genome/aspergillus_group/MultiHome.html) such as *Aspergillus* Comparative Database and the *Fusarium* Comparative Database have allowed the identification hundreds of annotated C6-encoding genes for each genus. A search of other genome databases at the Broad Institute for the latter three phyla found that *Ustilago maydis* (corn smut) and *Coprinopsis cinerea* of *Basidiomycota*, *Rhizopus oryzae* of *Mucoromycotina*, and *Allomyces macrogynus* and *Spizellomyces punctatus* of *Cytridiomycota* also contain annotated C6 zinc cluster proteins although fewer in number. Likely, C6 proteins are abundant in all fungal species. Over the past 25 years, the functions of only about 30 to 40 of ascomycete C6 proteins have been characterized. These proteins are primarily associated with regulation of genes involved in three classes of function: (1) utilization of carbon and nitrogen substrates/compounds, (2) production of secondary metabolites, and (3) asexual and sexual development.

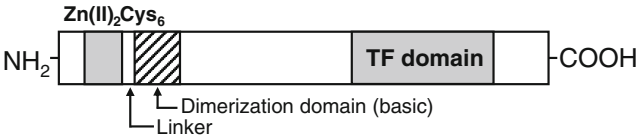

Fig. 1 Schematic representation of the general structure of fungal $Zn(II)_2Cys_6$ zinc cluster proteins. C6 proteins commonly contain two functional domains, the DNA-binding domain which includes the $Zn(II)_2Cys_6$ motif, the linker region and downstream basic dimerization region, and the regulatory domain which is a specific transcription factor (TF) domain

Fungal C6 proteins involved in regulation of carbon and nitrogen utilization

A majority of the fungal C6 proteins involved in the regulation of genes necessary for carbon and nitrogen utilization have been identified mainly from *A. nidulans*, a fungus long been used as a genetic and molecular model. These include AlcR in ethanol metabolism (Felenbok et al. 1988), FacB in acetate utilization (Todd et al. 1997), QutA in quinate utilization (Beri et al. 1987), AmdR in catabolism of acetamide and omega amino acids (Andrianopoulos and Hynes 1990), PrnA in proline utilization (Scazzocchio 1994), UaY in purine catabolism (Suarez et al. 1995), and NirA in nitrate assimilation (Burger et al. 1991). A few of these protein homologs also have been characterized in another model fungus *N. crassa*, such as ACU15 (FacB) (Bibbins et al. 2002), QA1F (QutA) (Baum et al. 1987), PCO1 (UaY) (Liu and Marzluf 2004), and NIT4 (NirA) (Yuan et al. 1991). Only one, HmgR, for tyrosine degradation has been characterized in the human pathogen *Aspergillus fumigatus* (Keller et al. 2011). C6 proteins that regulate genes involved in degradation of complex carbohydrates have been mainly reported for industrially important fungi, for example, AmyR of *Aspergillus oryzae* that regulates expression of clustered amylolytic genes of *agdA* (encoding alpha-glucosidase) and *amyA* (encoding alpha-amylase) (Gomi et al. 2000), XlnR of *A. oryzae* that regulates expression of more than 30 xylanolytic and cellulolytic genes in the degradation of beta-1,4-xylan, arabinoxylan, cellulose, and xyloglucan and catabolism of mono sugars (Noguchi et al. 2009), ManR of *A. oryzae* that regulates expression of the endo-ß-mannase gene (Ogawa et al. 2012) and InuR of *Aspergillus niger* that regulates expression of inulinolytic and sugar transport genes (Yuan et al. 2008).

Fungal C6 proteins involved in regulation of biosynthesis of secondary metabolites

Fungi are capable of producing many low molecular weight, structurally heterogeneous secondary metabolites. These

compounds are not required for growth of the producing fungus, and are, therefore, considered secondary metabolites. Some secondary metabolites known as mycotoxins are toxic to humans and animals, but many other secondary metabolites have important pharmacological applications (Brakhage 2012). The C6 proteins that regulate genes involved in secondary metabolite production function as transcription activators to upregulate expression of clustering genes. One of the best known examples is AflR of *A. flavus*, *Aspergillus parasiticus*, and *A. nidulans* AflR, which controls expression of pathway genes for the production of aflatoxin and sterigmatocystin, respectively (Brown et al. 1996; Chang et al. 1995; Payne et al. 1993). A few other C6 regulators involved in mycotoxin production include *Fusarium verticillioides* FUM21 for the biosynthesis of fumonisins, which cause leukoencephalomalacia in equids and pulmonary edema in swine (Brown et al. 2007), DEP6 of *Alternaria brassicicola* for the biosynthesis of depudecin, a histone deacetylase inhibitor (Wight et al. 2009), and GliZ of *A. fumigatus* for the biosynthesis of gliotoxin, an epipolythiodioxopiperazine metabolite and a virulence factor (Bok et al. 2006). SirZ of *Leptosphaeria maculans*, which is homologous to *A. fumigatus* GliZ, is required for biosynthesis of the phytotoxin, sirodesmin (Fox et al. 2008). In *Cercospora nicotianae*, CTB8 regulates genes required for the biosynthesis of the host non-selective photoactivated phytotoxin, cercosporin (Chen et al. 2007). C6 regulators also are required for the biosynthesis of several therapeutical agents. For example, LovE of *Aspergillus terreus* for the biosynthesis of the cholesterol-lowering compound, lovastatin (Huang and Li 2009). Two LovE homologs, MokH and MlcR, required for the biosynthesis of cholesterol-lowering metabolites, monacolin K and compactin, respectively, also have been studied in *Monascus pilosus* (Chen et al. 2010) and *Penicillium citrinum* (Abe et al. 2002). ApdR, AfoA, and MdpE of *A. nidulans* are required for the biosynthesis of anti-cancer compounds, aspyridones (Bergmann et al. 2007), asperfuranone (Chiang et al. 2009), and monodictyphenone (Chiang et al. 2010), respectively. However, CtnA of *Monascus purpureus*, a homolog of *A. nidulans* AfoA, is involved in the biosynthesis of the nephrotoxic polyketide citrinin (Shimizu et al. 2007). Pigments constitute another group of fungal secondary metabolites that have important functions, including infection of hosts and protection cells from photo damages. Cmr1p of *Colletotrichum lagenarium* regulates melanin biosynthesis as do its counterparts of Pig1p in *Magnaporthe grisea* (Tsuji et al. 2000) and BMR1 in *Bipolaris oryzae* (Kihara et al. 2008). Bik4 is required for biosynthesis of the red pigment bikaverin in *Fusarium fujikuroi* (Wiemann et al. 2009). GIP2 regulates biosynthesis of the mycelial pigment aurofusarin in *Gibberella zeae* (anamorph: *Fusarium graminearum*) (Kim et al. 2006).

Identification of additional genes encoding a $Zn(II)_2Cys_6$ domain in the *A. flavus* genome database

The *Aspergillus* Comparative Database at the Broad Institute contains many *A. flavus* genes annotated to encode $Zn(II)_2Cys_6$ proteins. When we performed a keyword search combined with the search option "find other genes with this domain" in 2010, 199 genes encoding proteins with a C6 domain were found. Eighty-two of the 199 genes also were annotated to encode a fungal-specific transcription factor (TF) domain. A similar keyword search with "fungal-specific transcription factor" showed that 200 were TF domain-encoding genes. Excluding the 82 genes annotated to encode proteins containing both a C6 and a TF domain, 117 were found to encode proteins only with a C6 domain and 118 only with a TF domain (Fig. 2). A recent examination of the updated *Aspergillus* Comparative Database indicates that 94 *A. flavus* genes encode only a C6 domain, 159 genes encode only a TF domain, and 96 genes encode both a C6 domain and a TF domain when duplicated records are removed. In an effort to obtain a more accurate count for C6 proteins, a refined search protocol was employed. A gene annotated to encode only a TF domain was translated in three reading frames with the DNAMAN software (Lynnon Soft, Vandreuil, QC, Canada), and the resulting amino acid sequences were examined manually for possible presence of a C6 domain. The search for a C6 domain took into consideration (1) the possible presence of intron(s) in the genomic DNA sequence assuming typical fungal intron sizes of 60 to 150 bp and (2) the expected location of a C6 domain relative to a TF domain (Fig. 1). In some cases, such as those in which a predicted TF domain is encoded by sequence near the 5′region proximal to the translational start site, an upstream nucleotide sequence region of 0.5 to 1.0 kb

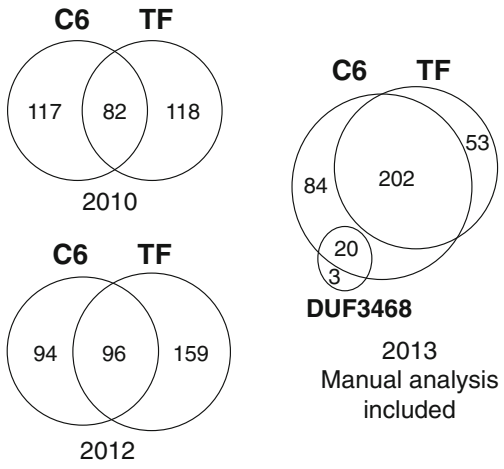

Fig. 2 Estimates of C6-encoding genes in the *A. flavus* genome annotated by automatic and manual analyses. *TF* fungal-specific transcription factor domain, *DUF3468* (Domain of Unknown Function)

was retrieved from the genome sequence database and included in the three-frame translational analysis. Among the 159 genes annotated to encode only a TF domain, an additional 106 genes were found to encode a C6 domain via this manual analysis (Supplemental Table S1). The 94 genes encoding only a C6 domain were translated in three reading frames, and the amino acid sequences were further analyzed by Conserved Domain (CD) search against the Conserved Domain Database (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). Ten were found to encode a domain called DUF3468 (DUF, domain of unknown function) by the CD search. Ten additional genes of this type, which encode both a DUF3468 and a C6 domain, were identified from the "DUF3468" keyword search of the *A. flavus* genome and subsequent manual analysis (see Domain of unknown function, DUF3468). No genes encoding both a TF domain and a DUF3468 were found. Adding these additional 116 putative C6-encoding genes to those currently annotated in the *Aspergillus* Comparative Database increases the total number from 190 to 306 (Supplemental Table S1), an increase of greater than 50 %. Figure 2 summarizes the estimated total number of C6-encoding genes from automatic annotation and manual identification along with data from the 2010 *A. flavus* genome database and the updated 2012 database. With a genome size of about 37 Mb, *A. flavus* is predicted to encode about 12,000 proteins (Payne et al. 2006). Therefore, about 2.5 % of the total predicted proteins of *A. flavus* are C6 proteins. This percentage is much higher than those calculated for *S. cerevisiae* (Goffeau et al. 1996) and *Candida albicans* (Maicas et al. 2005), which is about 0.9 and 1.2 %, respectively.

Average distribution of C6-encoding genes in *A. flavus*, *A. nidulans*, and yeast genomes

The numbers (in parenthesis) of C6 proteins for other *Aspergillus* species identified through automated annotation and listed in the *Aspergillus* Comparative Database are as follows: *A. clavatus* (180), *A. fumigatus* (186), *A. nidulans* (243), *A. niger* (236), *A. terreus* (181), and *A. oryzae* (177). The genome size of *A. flavus* is 36.8 Mb and *A. nidulans* 30.1 Mb. The estimated total numbers of the C6 proteins from the two aspergilli are comparable when taking into consideration the genome size of respective species. If distributed evenly, approximately one C6-encoding gene would reside in each 130-kb genomic region. All other species in the genus *Aspergillus* whose genomes have been sequenced have eight chromosomes, but their genome sizes vary. The genome size of *A. fumigatus* is 29.4 Mb, *A. niger* 37.2 Mb, *A. terreus* 29.3 Mb, *A. clavatus* 27.9 Mb, and *A. oryzae* 37.1 Mb. As for *A. flavus*, the total numbers of C6 factors for these aspergilli likely have been underestimated. Refinement of the gene-call algorithms and bioinformatic protocols will undoubtedly increase the number of C6-encoding genes identified significantly. In yeast, 54 and 70 C6-containing proteins have been reported for *S. cerevisiae* (Akache et al. 2001; MacPherson et al. 2006) and *C. albicans* (Maicas et al. 2005), which have a genome size of 11.8 and 14.5 Mb, respectively. The C6-encoding gene frequency in yeast genome is equivalent to one in 200 kb, which apparently is much lower than that estimated for aspergilli. Genome augmentation via duplication/acquisition in fungi probably is responsible for the marked increase in the number of C6 genes in order for the fungi to cope with a more complex environment and to occupy and adapt to specific living niches.

Sub-grouping of $Zn(II)_2Cys_6$ domains of *A. flavus* and *A. nidulans*

In the well-known Gal4 zinc cluster domain, C_1-X2-C_2-X6-C_3-X6-C_4-X2-C_5-X6-C_6, cysteine residues C_1, C_2, C_3 and C_4 are ligands for one zinc ion, and C_1, C_4, C_5 and C_6 are for another zinc ion (Gardner et al. 1991; Marmorstein et al. 1992; Pan and Coleman 1990). Thus, the first and fourth cysteine residues are shared by both zinc ions. Varying the number of residues between C_3 and C_4 and between C_5 and C_6 presumably relaxes the constraints resulting from the canonical subregions of C_1-C_2, C_2-C_3, and C_4-C_5 to allow better contact of both zinc ions with the cysteine residues. This, in turn, gives an optimal conformation for DNA recognition and binding. The number of amino acid residues in the C6 domains between C_1 and C_2 and between C_2 and C_3 in *A. flavus* and other fungal and yeast proteins is always 2 and 6, respectively. The variability in spacing between the other cysteines, C_3-C_4 and C_5-C_6, in the zinc cluster allows the *A. flavus* C6 domain proteins to be categorized into the subgroups shown in Table 1. Proteins having the pattern of C-2-C-6-C-6-C-2-C-6-C are most abundant followed by those having the pattern C-2-C-6-C-5-C-2-C-6-C. The ratio of predicted proteins with these patterns is about 2:1. Other less common patterns include C-2-C-6-C-5-C-2-C-8-C and C-2-C-6-C-6-C-2-C-8-C. A similar proportion of the classified subgroups are found for *A. nidulans*, which suggests that the functions of these variant C6 proteins among ascomycete fungi are evolutionarily conserved. All the C6 transcription factors characterized so far bind with sequence specificity to sites mainly consisting of GC-rich terminal trinucleotides that are separated by a variable internal space sequence. The terminal trinucleotides are usually palindromic but in some cases occur as direct repeats. Based on the limited numbers of C6 domains and recognition sites characterized (MacPherson et al. 2006; Todd and Andrianopoulos 1997), no correlation between the identified subgroups and the spacing

Table 1 Zinc cluster DNA-binding domains of *A. flavus* and *A. nidulans*

Subgroup	*A. flavus*	*A. nidulans*
C-2-C-6-C-5-C-2-C-6-C	87	70
C-2-C-6-C-6-C-2-C-6-C	139	128
C-2-C-6-C-7-C-2-C-6-C	8	8
C-2-C-6-C-8-C-2-C-6-C	16	25
C-2-C-6-C-9-C-2-C-6-C	12	9
C-2-C-6-C-10-C-2-C-6-C	5	4
C-2-C-6-C-12-C-2-C-6-C	2	2
C-2-C-6-C-15-C-2-C-6-C	1[a]	0
C-2-C-6-C-5-C-2-C-7-C	3	0
C-2-C-6-C-6-C-2-C-7-C	1	1
C-2-C-6-C-7-C-2-C-7-C	1	0
C-2-C-6-C-8-C-2-C-7-C	4	1
C-2-C-6-C-5-C-2-C-8-C	18	18
C-2-C-6-C-6-C-2-C-8-C	8	7
C-2-C-6-C-6-C-2-C-9-C	3	2
C-2-C-6-C-5-C-2-C-11-C	1	1

[a] *A. flavus alcR* (AFL2G_02974); *A. nidulans alcR* is ANID_08978.1 that encodes a domain of C-2-C-6-C-16-C-2-C-6-C.

of nucleotides in the binding sites has been identified. The linker region that is located carboxy-terminally to the C6 domain and positioned before the dimerization region likely also plays a significant role in mediating the sequence-specific C6 binding to DNA (Reece and Ptashne 1993).

Domain of unknown function, DUF3468

As mentioned earlier, among the 94 C6-encoding genes that were predicted by the Conserved Domain search not to encode a fungal specific TF domain, ten were found to encode a unique domain called DUF3468 (DUF, domain of unknown function) This domain is present in a family of putative fungal transcription factors typically at the carboxyl region with a size of 350 to 400 amino acids. A "DUF3468" keyword search of the *Aspergillus* Comparative Database revealed a total of 23 annotated DUF3468 proteins. Manual analyses of the remaining 13 DUF3468 proteins indicate that 10 additional proteins contain a C6 domain. The 20 genes are AFL2G_00121.2, AFL2G_00473.2, AFL2G_01202.2, A AFL2G_01693.2, FL2G_03094.2, AFL2G_03721.2, AFL2G_03753.2, AFL2G_04415.2, AFL2G_06402.2, AFL2G_06574.2, AFL2G_07853.2, AFL2G_07980.2, AFL2G_08040.2, AFL2G_08203.2, AFL2G_09406.2, AFL2G_09466.2, AFL2G_09728.2, AFL2G_09865.2, AFL2G_11881.2, and AFL2G_12301.2. All have the C6 pattern of C-2-C-6-C-6-C-2-C-6-C. The remaining three that do not encode a C6 domain are AFL2G_00885.2, AFL2G_05084.2, and AFL2G_08434.2. Other genomes of aspergilli in the *Aspergillus* Comparative Database also contain various numbers of genes encoding proteins with a DUF3468 domain.

C6 regulators with DUF3468 involved in asexual conidiation of *A. nidulans* and *A. flavus*

In *A. nidulans*, two C6-encoding genes, *oefC* (overexpressed fluffy, AY792357) and *sfgA* (suppressor of *fluG*, DQ087435), involved in asexual development have been characterized. Lee et al. (2005) introduced into the *A. nidulans* wild-type strain genomic library clones that were placed under the control of nitrite reductase (*niiA*) gene promoter. They cloned the *oefC* gene that rendered transformants to produce fluffy, undifferentiated aerial hyphae and failed to develop conidiophores under stress conditions that induce asexual conidiation (Lee et al. 2005). In *A. nidulans*, mutations in the *fluG* gene abolished the induction of conidiation and also resulted in cotton-like fluffy colonies. Overexpression of the full-length *fluG* or the portion encoding the C-terminal half portion caused abnormal conidiophore development in liquid submerged culture that suppresses conidiation (D'Souza et al. 2001; Lee and Adams 1996). Seo et al. (2006) found that impairment in *sfgA* can suppress the conidiation defect in the *A. nidulans* Δ*fluG* mutant. Since deletion of *sfgA* bypassed the need for *fluG* in conidiation and overexpression of *sfgA* inhibited conidiation, they proposed SfgA as a repressor that interacts with FluG (Seo et al. 2006). Our analyses indicate that *A. nidulans* OefC and SfgA both contain a DUF3468 domain of about 380 and 420 amino acid residues, respectively, but the two DUF3468 domains have only 17 % amino acid sequence identity. The orthologs of *oefC* and *sfgA* in *A. flavus* (AFL2G_01202.2 and AFL2G_09865.2) also encode C6 and DUF3468 domain proteins. The DUF3468 domains of *A. flavus* OefC and SfgA share 82 and 74 % identity to those of *A. nidulans*, respectively (Fig. 3). This suggests that the DUF3468 domain may be engaged in specific interactions with known proteins that control development such as FluG (Chang et al. 2012) or one or more components in the *velvet* complex, VelB/VeA/LaeA (Bayram et al. 2008).

Presence of DUF3468 in C6 regulators for *A. nidulans* sexual differentiation

Two *A. nidulans* C6-encoding genes shown to be involved in sexual development, *rosA* (repressor of sexual development, AJ519682, ANID_05170.1) and *nosA* (number of sexual spores, AM231027, ANID_01848.1) (Vienken and Fischer 2006; Vienken et al. 2005), also encode proteins that possess C-terminal DUF3468 (Pfam: PF11951) domains

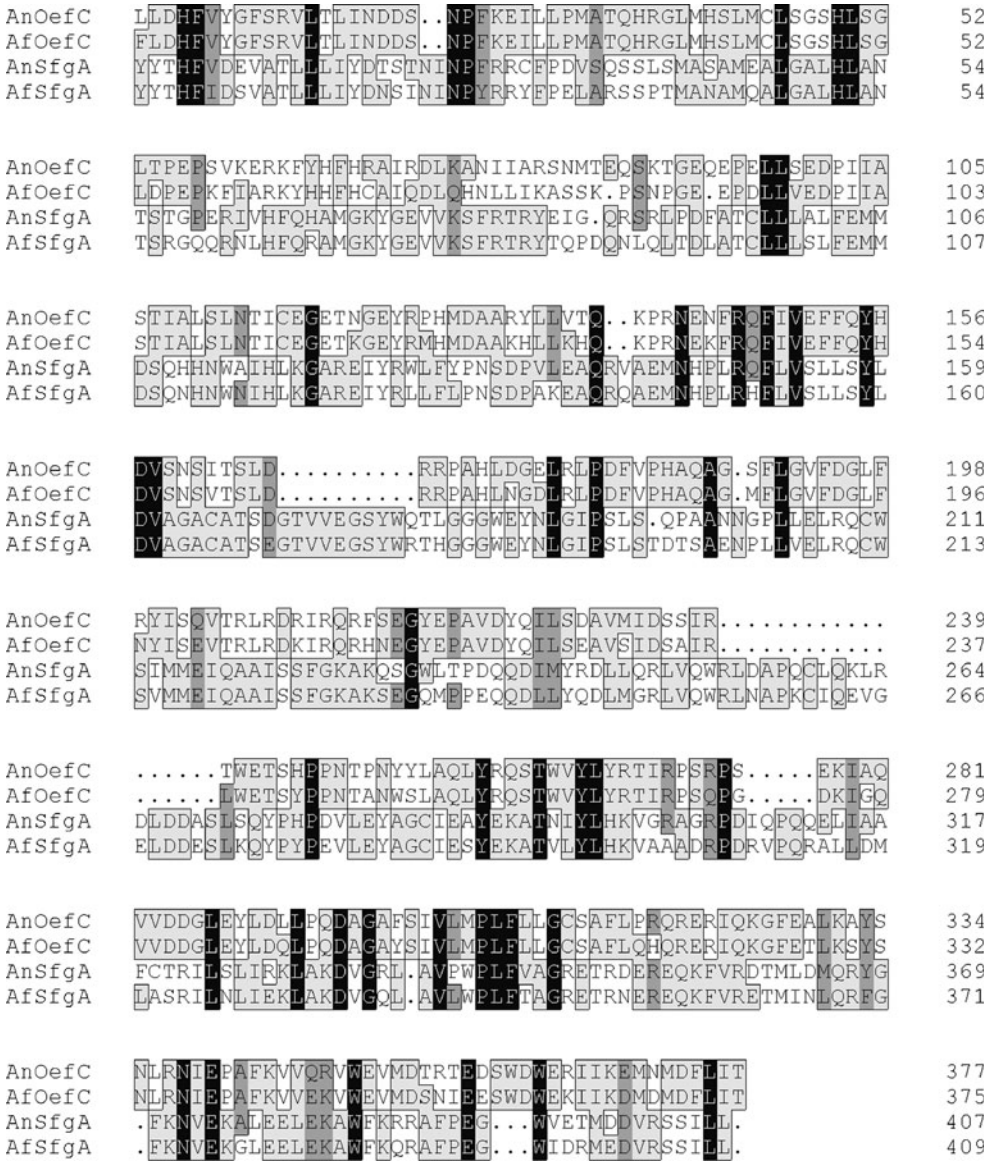

Fig. 3 Alignment of amino acid sequences of DUF3468 domains in OefC and SfgA of *A. nidulans* and *A. flavus*. DUF3468 identified by the NCBI Conserved Domain search is located in *A. nidulans* OefC (AAW55628) from 261 to 636, *A. nidulans* SfgA (AAY99779) from 194 to 600, *A. flavus* OefC from 261 to 635, and *A. flavus* SfgA from 166 to 574

that were revealed by our Conserved Domain search. These two DUF3468 domains share 51 % amino acid sequence identity. *A. nidulans* RosA downregulates expression of the sexual development regulatory genes *nsdD*, *veA*, and *stuA*. Overexpression of *rosA* resulted in colonies with fluffy cotton-like hyphae (Vienken et al. 2005). The *A. nidulans nosA* gene, upregulated during the late asexual development, is required for the completion of the sexual cycle. Defects in *nosA* block at the primordial stage but occasionally produced minute cleistothecia containing fertile ascospores (Vienken and Fischer 2006). AFL2G_01801.2 of *A. flavus* is the ortholog of *A. nidulans nosA* with 71 % identity and 82 % positive between predicted amino acid sequences. *A. flavus* NosA also are C6 proteins with a DUF3468 domain. AFL2G_03812. 2 is the ortholog of *A. nidulans rosA* (48 % identity and 64 % positive) and the encoded RosA has a DUF3468 domain. However, the region corresponding to the *A. nidulans* RosA C6-containing portion has been replaced by a PAT1/TFIIA/DUF1421 domain. Being heterothallic and reproducing largely through asexual conidiation, *A. flavus* appears to have lost some of the components involved in homothallic sexual development. Although sexual reproduction under laboratory conditions has recently been demonstrated with *A. flavus* strains of different mating types (Horn et al. 2009), strict regulation on the sexual cycle may no longer be necessary for *A. flavus*.

Physical association of C6-encoding genes with *A. flavus* secondary metabolite gene clusters

A. flavus is known to produce as many as 27 secondary metabolites (Pildain et al. 2008). These include the well-

Table 2 Physical association of C6 domain genes with the 55 gene clusters in *A. flavus* genome

Cluster	Backbone gene	KEGG locus	ACD locus	$Zn(II)_2Cys_6$ Gene ID	Metabolite relationship[a]
1	PKS	AFLA_002900	AFL2G_09607.2		
2	DMTS	AFLA_004300	AFL2G_09741.2	AFLA_004280[b]	
3	NRPS	AFLA_004450	AFL2G_09757.2		
4	NRPS	AFLA_005440	AFL2G_09859.2	AFL2G_09865.2	
5	PKS	AFLA_006170	AFL2G_09923.2		Conidial pigment
6	NRPS	AFLA_008770	AFL2G_12042.2		
7	NRPS-like	AFLA_009120	AFL2G_12077.2		
7	PKS	AFLA_009140	AFL2G_12079.2		
8	PKS	AFLA_010000	AFL2G_12161.2		
8	NRPS	AFLA_010010	AFL2G_12161.2		
8	NRPS	AFLA_010020	AFL2G_12162.2		
9	NRPS	AFLA_010580	AFL2G_12207.2		Siderophore
9	NRPS	AFLA_010620	AFL2G_12211.2		
10	Scytalone dehydratase	AFLA_016140	AFL2G_03259.2		
11	NRPS-like	AFLA_023020	AFL2G_01550.2	AFL2G_01551.2	
12	NRPS-like	AFLA_028720	AFL2G_02082.2		
13	NRPS	AFLA_038600	AFL2G_04847.2		
14	IroE-like	AFLA_041050	AFL2G_05061.2	AFL2G_05056.2	Siderophore
15	DMTS	AFLA_045490	AFL2G_05466.2	AFL2G_05459.2	Aflatrem, *ATM2*
16	PKS-like	AFLA_053770	AFL2G_10571.2	AFL2G_10570.2	
16	PKS-like	AFLA_053780	AFL2G_10571.2		
16	PKS	AFLA_053870	AFL2G_10577.2		
17	NRPS-like	AFLA_054270	AFL2G_10612.2	AFL2G_10615.2	
18	PKS-like	AFLA_060010	AFL2G_06151.2	AFL2G_06146.2	
18	PKS-like	AFLA_060020	AFL2G_06151.2		
19	NRPS	AFLA_060680	AFL2G_06212.2		
20	PKS	AFLA_062820	AFL2G_06390.2		
20	PKS	AFLA_062860	AFL2G_06393.2		
21	NRPS	AFLA_064240	AFL2G_07262.2		Gliotoxin-like
21	NRPS	AFLA_064560	AFL2G_07288.2		
22	NRPS	AFLA_066720	AFL2G_07493.2	AFL2G_07485.2	
23	NRPS-PKS	AFLA_066840	AFL2G_07507.2	AFL2G_07508.2	
23	PKS	AFLA_066980	AFL2G_07518.2	AFL2G_07511.2	
24	NRPS	AFLA_069330	AFL2G_07731.2		
25	IPNS	AFLA_070870	AFL2G_07876.2		Penicillin-like
25	NRPS-like	AFLA_070920	AFL2G_07881.2	AFL2G_07886.2 AFL2G_07887.2	
26	PKS-like	AFLA_079360	AFL2G_00677.2	AFL2G_00673.2	
26	NRPS-like	AFLA_079380	AFL2G_00677.2		
26	NRPS-like	AFLA_079400	AFL2G_00680.2		
27	PKS	AFLA_082150	AFL2G_00935.2	AFL2G_00934.2	Sclerotial pigment
28	NRPS-like	AFLA_082480	AFL2G_00966.2	AFL2G_00969.2	
29	DMTS	AFLA_084080	AFL2G_01107.2	AFL2G_01108.2	
30	DMTS	AFLA_090190	AFL2G_08061.2		
30	NRPS	AFLA_090200	AFL2G_08062.2	AFL2G_08058.2	
31	NRPS-like	AFLA_095040	AFL2G_08520.2		
32	GGPPS	AFLA_096390	AFL2G_08643.2	AFL2G_08641.2	Aflatrem, *ATM1*
33	NRPS-like	AFLA_096700	AFL2G_08672.2		
33	NRPS-like	AFLA_096710	AFL2G_08672.2	AFL2G_08674.2	
33	PKS	AFLA_096770	AFL2G_08678.2		Lovastatin-like

Table 2 (continued)

Cluster	Backbone gene	KEGG locus	ACD locus	$Zn(II)_2Cys_6$ Gene ID	Metabolite relationship[a]
34	NRPS	AFLA_100340	AFL2G_10935.2	AFLA_100300[c]	
35	NRPS-like	AFLA_101700	AFL2G_11054.2	AFL2G_11045.2	
36	PKS-like	AFLA_104210	AFL2G_03890.2	AFL2G_03891.2	Phenolphthiocerol-like
36	PKS-like	AFLA_104240	AFL2G_03893.2		
36	PKS-like	AFLA_104250	AFL2G_03894.2		
37	NRPS-like	AFLA_105190	AFL2G_03983.2	AFL2G_03975.2	
38	PKS	AFLA_105450	AFL2G_04006.2	AFL2G_04013.2	
39	PKS	AFLA_108550	AFL2G_04285.2		
40	PKS	AFLA_112840	AFL2G_04689.2	AFL2G_04688.2[d]	
41	PKS	AFLA_114820	AFL2G_12403.2		6-MSA-like
42	PKS	AFLA_116220	AFL2G_11312.2	AFL2G_11313.2	
43	PKS-like	AFLA_116500	AFL2G_11338.2		
43	DMTS	AFLA_116600	AFL2G_11348.2	AFL2G_11355.2	
44	PKS	AFLA_116890	AFL2G_11372.2	AFL2G_11371.2	
45	NRPS-like	AFLA_118440	AFL2G_11528.2		
46	PKS	AFLA_118940	AFL2G_11574.2		
46	PKS	AFLA_118960	AFL2G_11576.2		
47	NRPS-like	AFLA_119110	AFL2G_11593.2	AFL2G_11610.2	
48	NRPS-like	AFLA_121520	AFL2G_11806.2	AFLA_121620[e]	
49	PKS-like	AFLA_125630	AFL2G_08911.2	AFL2G_08907.2	
49	PKS-like	AFLA_125640	AFL2G_08911.2		
50	PKS	AFLA_126710	AFL2G_09015.2		Fumonisin-like
51	PKS	AFLA_127090	AFL2G_09054.2	AFL2G_09045.2	Citrinin-like
52	PKS	AFLA_128060	AFL2G_09150.2	AFL2G_09159.2 AFL2G_09160.2	
53	NRPS	AFLA_135490	AFL2G_06882.2		
54	PKS	AFLA_139410	AFL2G_07228.2	AFL2G_07224.2	Aflatoxin
55	DMTS	AFLA_139480	AFL2G_07235.2		
55	NRPS-PKS	AFLA_139490	AFL2G_07236.2	AFL2G_07237.2	Cyclopiazonic acid

PKS polyketide synthase, *DMTS* dimethylallyl tryptophan synthase, *NRPS* nonribosomal peptide synthase, *IroE* putative enterobactin esterase, *IPNS* isopenicillin N synthase, *GGPPS* geranylgeranyl pyrophosphate synthase, *KEGG* Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp/kegg/), *ACD Aspergillus* Comparative Database of the Broad Institute, *6-MSA* 6-methyl salicyclic acid

[a] Production of aflatoxin, aflatrem, cyclopiazonic acid, conidial pigment, and sclerotial pigment has been experimentally confirmed. Other metabolites are inferred from putative identities of the related backbone enzymes

[b] Located between AFL2G_9737 and AFL2G_9738 but not annotated in ACD

[c] The approximate locus is AFL2G_10931.2

[d] AFL2G_04688.2 encodes a unique C6 pattern of 2-6-4-2-6

[e] AFLA_121620=CQACVRGKRRCDQLWPRCSRCQARGIEC; No match found in ACD

characterized toxic compounds: aflatoxin, cyclopiazonic acid, and aflatrem as well as metabolites involved in conidial and sclerotial pigmentation, and melanin formation. One significant hallmark of fungal genes involved in secondary metabolite biosynthesis pathways is that these genes are usually found in individual clusters. Expression of biosynthesis genes in a few known clusters is co-regulated by a C6 transcription factor encoded by the regulatory gene located in the same gene cluster. The program "Secondary Metabolite Unknown Regions Finder" (SMURF; http://www.jcvi.org/smurf) was developed to predict gene clusters in fungal genomes (Khaldi et al. 2010). SMURF searches for the so-called backbone genes that encode multifunctional enzymes associated with production of four classes of secondary metabolites. The backbone genes are those encoding polyketide synthase (PKS) for polyketide, nonribosomal peptide synthetase (NRPS) for nonribosomal peptide, NRPS–PKS for a hybrid metabolite, and prenyltransferase for terpenoid. After a backbone gene is located, SMURF then analyzes neighboring genes for encoded canoncial domains, such as those found in reductive and oxidative enzymes and methyltransferases that are commonly associated

with further modifications of the metabolite formed by the backbone enzyme. SMURF analysis of the *A. flavus* genome sequence has predicted 55 secondary metabolite gene clusters (Georgianna et al. 2010; Table 2). Studies have confirmed that clusters 5, 54, and 55 are involved in biosynthesis of conidial pigment (Chang et al. 2010), aflatoxins (Yu et al. 2004), and cyclopiazonic acid (Chang et al. 2009), respectively. The aflatrem biosynthesis gene cluster is split into two loci; the first locus, *ATM1*, is telomere proximal on chromosome 5 and contains three genes, and the second locus, *ATM2*, is telomere distal on chromosome 7 and contains five genes (Nicholson et al. 2009; Zhang et al. 2004). *ATM1* corresponds to cluster 32 and *ATM2* corresponds to cluster 15. *ATM2* contains a C6-encoding gene and *ATM1* has one adjacent to it. Biosynthesis genes of *ATM1* and *ATM2* are able to complement *Penicillium paxilli* deletion mutants defective in biosynthesis of paxilline, an indole-diterpene tremorgen (Nicholson et al. 2009; Young et al. 2001), but no studies have confirmed the function of the two C6-encoding genes physically associated with *ATM1* and *ATM2*. In the closely related *A. oryzae*, overexpression of the *aoiH* (=AFLA_116230) C6-encoding gene in a gene cluster that is equivalent to *A. flavus* gene cluster 42 activates a silent biosynthetic pathway to produce a novel polyketide metabolite (Nakazawa et al. 2012). The *aoiH* homologue, AFL2G_11313.2, likely is a pathway-specific regulatory gene of *A. flavus* cluster 42. Evidence also has been obtained for AFL2G_00934.2 to be the regulatory gene of cluster 27 which encodes proteins involved in biosynthesis of a sclerotial pigment (Cary, personal communication). The metabolites produced by most of the other clusters are largely unknown. In gene clusters of 27 and 42, a C6-encoding gene is located right next to the respective polyketide synthase gene. The *aflR* gene (AFL2G_07224.2) of cluster 54 required for aflatoxin biosynthesis is adjacent to the polyketide synthase gene, *pksA* (Yu et al. 2004). Whether or not a close physical association of a C6-encoding gene with a backbone gene(s) is indicative of its functional involvement is still not clear. With the identification of the majority of C6-encoding genes from the *A. flavus* genome, an effort was made to assign specific C6-encoding genes within a span of ten-gene distance to the 55 clusters. Approximately half of the 55 gene clusters are associated with a C6-encoding gene (Table 2). Some clusters like 25 and 52 have two C6-encoding genes adjacent to the backbone gene. It is possible that not all cluster-associated C6-encoding genes regulate expression of adjacent clustering genes. For example, cluster 55 is involved in cyclopiazonic acid (CPA) biosynthesis, but disruption of AFL2G_07237.2 in the gene cluster did not affect CPA production (Chang et al. 2009). The observed frequent association may be in part due to the high number of C6 domain-encoding genes (2.5 %) in the *A. flavus* genome.

Future perspectives

With increasing numbers of fungal genomes being sequenced, a wealth of information concerning gene sequence and location is becoming readily available. Bioinformatics has expanded our ability to predict gene function and analyze organization of gene clusters. Comparative genome studies have been performed to decipher evolutionary relationship among related species (Galagan et al. 2005; Payne et al. 2006; Sato et al. 2011) or among strains of the same species (Borneman et al. 2011). Emphasis now must be shifted toward examining functions of annotated groups of genes. Current protocols for automatic gene prediction are still far from perfect. Refinement of bioinformatic algorithms to enhance accuracy of gene prediction and annotation therefore is a prerequisite for the advance of functional genomics studies. Comparison of C6 domains and the normally conserved downstream basic amino acid dimerization region will spur investigation of mechanisms of phylogenetic diversity among different fungal species. The central role played by the C6 proteins has been evident in either as activators or repressors to modulated expression of controlled genes. Further understanding of how C6-encoding genes are activated and how C6 proteins are posttranslationally modified and interact with co-activators or globally acting transcription factors via the TF or DUF3468 domain needs to be pursued. Their roles in basic fungal development and differentiation also are largely unknown. Association of abilities to infect and colonize host plants with C6 proteins (Bluhm et al. 2008; Imazaki et al. 2007) is another new but rarely explored field. C6 proteins have been implicated in multidrug resistance and in response to stress such as heat shock, low pH, and high osmolarity in *S. cerevisiae* (Akache et al. 2001; MacPherson et al. 2006). However, no studies have probed this important area of transcription regulation which is critical for fungal survival. Challenges and surprises will arise by future studies of this fundamental class of regulators.

References

Abe Y, Ono C, Hosobuchi M, Yoshikawa H (2002) Functional analysis of *mlcR*, a regulatory gene for ML-236B (compactin) biosynthesis in *Penicillium citrinum*. Mol Genet Genomics 268:352–361

Akache B, Wu K, Turcotte B (2001) Phenotypic analysis of genes encoding yeast zinc cluster proteins. Nucleic Acids Res 29:2181–2190

Andrianopoulos A, Hynes MJ (1990) Sequence and functional analysis of the positively acting regulatory gene *amdR* from *Aspergillus nidulans*. Mol Cell Biol 10:3194–3203

Baum JA, Geever R, Giles NH (1987) Expression of *qa-1F* activator protein: identification of upstream binding sites in the *qa* gene

cluster and localization of the DNA-binding domain. Mol Cell Biol 7:1256–1266

Bayram O, Krappmann S, Ni M, Bok JW, Helmstaedt K, Valerius O, Braus-Stromeyer S, Kwon NJ, Keller NP, Yu JH, Braus GH (2008) VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. Science 320:1504–1506

Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA, Hertweck C (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. Nat Chem Biol 3:213–217

Beri RK, Whittington H, Roberts CF, Hawkins AR (1987) Isolation and characterization of the positively acting regulatory gene *QUTA* from *Aspergillus nidulans*. Nucleic Acids Res 15:7991–8001

Bibbins M, Crepin VF, Cummings NJ, Mizote T, Baker K, Mellits KH, Connerton IF (2002) A regulator gene for acetate utilisation from *Neurospora crassa*. Mol Genet Genomics 267:498–505

Bluhm BH, Kim H, Butchko RA, Woloshuk CP (2008) Involvement of ZFR1 of *Fusarium verticillioides* in kernel colonization and the regulation of *FST1*, a putative sugar transporter gene required for fumonisin biosynthesis on maize kernels. Mol Plant Pathol 9:203–211

Bok JW, Chung D, Balajee SA, Marr KA, Andes D, Nielsen KF, Frisvad JC, Kirby KA, Keller NP (2006) GliZ, a transcriptional regulator of gliotoxin biosynthesis, contributes to *Aspergillus fumigatus* virulence. Infect Immun 74:6761–6768

Borneman AR, Desany BA, Riches D, Affourtit JP, Forgan AH, Pretorius IS, Egholm M, Chambers PJ (2011) Whole-genome comparison reveals novel genetic elements that characterize the genome of industrial strains of *Saccharomyces cerevisiae*. PLoS Genet 7:e1001287

Brakhage AA (2012) Regulation of fungal secondary metabolism. Nat Rev Microbiol 11:21–32

Brown DW, Butchko RA, Busman M, Proctor RH (2007) The *Fusarium verticillioides* FUM gene cluster encodes a $Zn(II)_2Cys_6$ protein that affects FUM gene expression and fumonisin production. Eukaryot Cell 6:1210–1218

Brown DW, Yu JH, Kelkar HS, Fernandes M, Nesbitt TC, Keller NP, Adams TH, Leonard TJ (1996) Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. Proc Natl Acad Sci U S A 93:1418–1422

Burger G, Strauss J, Scazzocchio C, Lang BF (1991) *nirA*, the pathway-specific regulatory gene of nitrate assimilation in *Aspergillus nidulans*, encodes a putative GAL4-type zinc finger protein and contains four introns in highly conserved regions. Mol Cell Biol 11:5746–5755

Chang P-K, Ehrlich KC, Yu J, Bhatnagar D, Cleveland TE (1995) Increased expression of *Aspergillus parasiticus aflR*, encoding a sequence-specific DNA-binding protein, relieves nitrate inhibition of aflatoxin biosynthesis. Appl Environ Microbiol 61:2372–2377

Chang P-K, Horn BW, Dorner JW (2009) Clustered genes involved in cyclopiazonic acid production are next to the aflatoxin biosynthesis gene cluster in *Aspergillus flavus*. Fungal Genet Biol 46:176–182

Chang P-K, Scharfenstein LL, Mack B, Ehrlich KC (2012) Deletion of the *Aspergillus flavus* orthologue of *A. nidulans fluG* reduces conidiation and promotes production of sclerotia but does not abolish aflatoxin biosynthesis. Appl Environ Microbiol 78:7557–7563

Chang P-K, Scharfenstein LL, Wei Q, Bhatnagar D (2010) Development and refinement of a high-efficiency gene-targeting system for *Aspergillus flavus*. J Microbiol Methods 81:240–246

Chen H, Lee MH, Daub ME, Chung KR (2007) Molecular analysis of the cercosporin biosynthetic gene cluster in *Cercospora nicotianae*. Mol Microbiol 64:755–770

Chen YP, Yuan GF, Hsieh SY, Lin YS, Wang WY, Liaw LL, Tseng CP (2010) Identification of the *mokH* gene encoding transcription factor for the upregulation of monacolin K biosynthesis in *Monascus pilosus*. J Agric Food Chem 58:287–293

Chiang YM, Szewczyk E, Davidson AD, Entwistle R, Keller NP, Wang CC, Oakley BR (2010) Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. Appl Environ Microbiol 76:2067–2074

Chiang YM, Szewczyk E, Davidson AD, Keller N, Oakley BR, Wang CC (2009) A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. J Am Chem Soc 131:2965–2970

D'Souza CA, Lee BN, Adams TH (2001) Characterization of the role of the FluG protein in asexual development of *Aspergillus nidulans*. Genetics 158:1027–1036

Endo H, Kajiwara S, Tsunoka O, Shishido K (1994) A novel cDNA, *priBc*, encoding a protein with a $Zn(II)_2Cys_6$ zinc cluster DNA-binding motif, derived from the basidiomycete *Lentinus edodes*. Gene 139:117–121

Felenbok B, Sequeval D, Mathieu M, Sibley S, Gwynne DI, Davies RW (1988) The ethanol regulon in *Aspergillus nidulans*: characterization and sequence of the positive regulatory gene *alcR*. Gene 73:385–396

Fox EM, Gardiner DM, Keller NP, Howlett BJ (2008) A $Zn(II)_2Cys_6$ DNA binding protein regulates the sirodesmin PL biosynthetic gene cluster in *Leptosphaeria maculans*. Fungal Genet Biol 45:671–682

Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR, Batzoglou S, Lee SI, Basturkmen M, Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scazzocchio C, Farman M, Butler J, Purcell S, Harris S, Braus GH, Draht O, Busch S, D'Enfert C, Bouchier C, Goldman GH, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, Selker EU, Archer DB, Penalva MA, Oakley BR, Momany M, Tanaka T, Kumagai T, Asai K, Machida M, Nierman WC, Denning DW, Caddick M, Hynes M, Paoletti M, Fischer R, Miller B, Dyer P, Sachs MS, Osmani SA, Birren BW (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. Nature 438:1105–1115

Gardner KH, Pan T, Narula S, Rivera E, Coleman JE (1991) Structure of the binuclear metal-binding site in the GAL4 transcription factor. Biochemistry 30:11292–11302

Georgianna DR, Fedorova ND, Burroughs JL, Dolezal AL, Bok JW, Horowitz-Brown S, Woloshuk CP, Yu J, Keller NP, Payne GA (2010) Beyond aflatoxin: four distinct expression patterns and functional roles associated with *Aspergillus flavus* secondary metabolism gene clusters. Mol Plant Pathol 11:213–226

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. Science 274:563–567

Gomi K, Akeno T, Minetoki T, Ozeki K, Kumagai C, Okazaki N, Iimura Y (2000) Molecular cloning and characterization of a transcriptional activator gene, *amyR*, involved in the amylolytic gene expression in *Aspergillus oryzae*. Biosci Biotechnol Biochem 64:816–827

Hidalgo P, Ansari AZ, Schmidt P, Hare B, Simkovich N, Farrell S, Shin EJ, Ptashne M, Wagner G (2001) Recruitment of the transcriptional machinery through GAL11P: structure and interactions of the GAL4 dimerization domain. Genes Dev 15:1007–1020

Hong M, Fitzgerald MX, Harper S, Luo C, Speicher DW, Marmorstein R (2008) Structural basis for dimerization in DNA recognition by Gal4. Structure 16:1019–1026

Horn BW, Moore GG, Carbone I (2009) Sexual reproduction in *Aspergillus flavus*. Mycologia 101:423–429

Huang X, Li HM (2009) Cloning and bioinformatic analysis of lovastatin biosynthesis regulatory gene *lovE*. Chin Med J (Engl) 122:1800–1805

Imazaki I, Kurahashi M, Iida Y, Tsuge T (2007) Fow2, a $Zn(II)_2Cys_6$-type transcription regulator, controls plant infection of the vascular wilt fungus *Fusarium oxysporum*. Mol Microbiol 63:737–753

Johnston M (1987) A model fungal gene regulatory mechanism: the GAL genes of *Saccharomyces cerevisiae*. Microbiol Rev 51:458–476

Keegan L, Gill G, Ptashne M (1986) Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science 231:699–704

Keller S, Macheleidt J, Scherlach K, Schmaler-Ripcke J, Jacobsen ID, Heinekamp T, Brakhage AA (2011) Pyomelanin formation in *Aspergillus fumigatus* requires HmgX and the transcriptional activator HmgR but is dispensable for virulence. PLoS One 6: e26604

Khaldi N, Seifuddin FT, Turner G, Haft D, Nierman WC, Wolfe KH, Fedorova ND (2010) SMURF: genomic mapping of fungal secondary metabolite clusters. Fungal Genet Biol 47:736–741

Kihara J, Moriwaki A, Tanaka N, Tanaka C, Ueno M, Arase S (2008) Characterization of the *BMR1* gene encoding a transcription factor for melanin biosynthesis genes in the phytopathogenic fungus *Bipolaris oryzae*. FEMS Microbiol Lett 281:221–227

Kim JE, Jin J, Kim H, Kim JC, Yun SH, Lee YW (2006) GIP2, a putative transcription factor that regulates the aurofusarin biosynthetic gene cluster in *Gibberella zeae*. Appl Environ Microbiol 72:1645–1652

Lee BN, Adams TH (1996) *fluG* and *flbA* function interdependently to initiate conidiophore development in *Aspergillus nidulans* through *brlA*β activation. EMBO J 15:299–309

Lee BY, Han SY, Choi HG, Kim JH, Han KH, Han DM (2005) Screening of growth- or development-related genes by using genomic library with inducible promoter in *Aspergillus nidulans*. J Microbiol 43:523–528

Liu TD, Marzluf GA (2004) Characterization of *pco-1*, a newly identified gene which regulates purine catabolism in *Neurospora*. Curr Genet 46:213–227

Ma J, Ptashne M (1987a) The carboxy-terminal 30 amino acids of GAL4 are recognized by GAL80. Cell 50:137–142

Ma J, Ptashne M (1987b) Deletion analysis of GAL4 defines two transcriptional activating segments. Cell 48:847–853

MacPherson S, Larochelle M, Turcotte B (2006) A fungal family of transcriptional regulators: the zinc cluster proteins. Microbiol Mol Biol Rev 70:583–604

Maicas S, Moreno I, Nieto A, Gomez M, Sentandreu R, Valentin E (2005) In silico analysis for transcription factors with $Zn(II)_2C_6$ binuclear cluster DNA-binding domains in *Candida albicans*. Comp Funct Genomics 6:345–356

Marmorstein R, Carey M, Ptashne M, Harrison SC (1992) DNA recognition by GAL4: structure of a protein–DNA complex. Nature 356:408–414

Nakazawa T, Ishiuchi K, Praseuth A, Noguchi H, Hotta K, Watanabe K (2012) Overexpressing transcriptional regulator in *Aspergillus oryzae* activates a silent biosynthetic pathway to produce a novel polyketide. ChemBioChem 13:855–861

Nicholson MJ, Koulman A, Monahan BJ, Pritchard BL, Payne GA, Scott B (2009) Identification of two aflatrem biosynthesis gene loci in *Aspergillus flavus* and metabolic engineering of *Penicillium paxilli* to elucidate their function. Appl Environ Microbiol 75:7469–7481

Noguchi Y, Sano M, Kanamaru K, Ko T, Takeuchi M, Kato M, Kobayashi T (2009) Genes regulated by AoXlnR, the xylanolytic and cellulolytic transcriptional regulator, in *Aspergillus oryzae*. Appl Microbiol Biotechnol 85:141–154

Ogawa M, Kobayashi T, Koyama Y (2012) ManR, a novel $Zn(II)_2Cys_6$ transcriptional activator, controls the beta-mannan utilization system in *Aspergillus oryzae*. Fungal Genet Biol 49:987–995

Pan T, Coleman JE (1990) GAL4 transcription factor is not a "zinc finger" but forms a $Zn(II)_2Cys_6$ binuclear cluster. Proc Natl Acad Sci U S A 87:2077–2081

Payne GA, Nierman WC, Wortman JR, Pritchard BL, Brown D, Dean RA, Bhatnagar D, Cleveland TE, Machida M, Yu J (2006) Whole genome comparison of *Aspergillus flavus* and *A. oryzae*. Med Mycol 44(Suppl):9–11

Payne GA, Nystrom GJ, Bhatnagar D, Cleveland TE, Woloshuk CP (1993) Cloning of the *afl-2* gene involved in aflatoxin biosynthesis from *Aspergillus flavus*. Appl Environ Microbiol 59:156–162

Pildain MB, Frisvad JC, Vaamonde G, Cabral D, Varga J, Samson RA (2008) Two novel aflatoxin-producing *Aspergillus* species from Argentinean peanuts. Int J Syst Evol Microbiol 58:725–735

Reece RJ, Ptashne M (1993) Determinants of binding-site specificity among yeast C6 zinc cluster proteins. Science 261:909–911

Sato A, Oshima K, Noguchi H, Ogawa M, Takahashi T, Oguma T, Koyama Y, Itoh T, Hattori M, Hanya Y (2011) Draft genome sequencing and comparative analysis of *Aspergillus sojae* NBRC4239. DNA Res 18:165–176

Scazzocchio C (1994) The proline utilisation pathway, history and beyond. Prog Ind Microbiol 29:259–277

Seo JA, Guan Y, Yu JH (2006) FluG-dependent asexual development in *Aspergillus nidulans* occurs via derepression. Genetics 172:1535–1544

Shimizu T, Kinoshita H, Nihira T (2007) Identification and in vivo functional analysis by gene disruption of *ctnA*, an activator gene involved in citrinin biosynthesis in *Monascus purpureus*. Appl Environ Microbiol 73:5097–5103

Suarez T, de Queiroz MV, Oestreicher N, Scazzocchio C (1995) The sequence and binding specificity of UaY, the specific regulator of the purine utilization pathway in *Aspergillus nidulans*, suggest an evolutionary relationship with the PPR1 protein of *Saccharomyces cerevisiae*. EMBO J 14:1453–1467

Todd RB, Andrianopoulos A (1997) Evolution of a fungal regulatory gene family: the $Zn(II)_2Cys_6$ binuclear cluster DNA binding motif. Fungal Genet Biol 21:388–405

Todd RB, Murphy RL, Martin HM, Sharp JA, Davis MA, Katz ME, Hynes MJ (1997) The acetate regulatory gene *facB* of *Aspergillus nidulans* encodes a $Zn(II)_2Cys_6$ transcriptional activator. Mol Gen Genet 254:495–504

Tsuji G, Kenmochi Y, Takano Y, Sweigard J, Farrall L, Furusawa I, Horino O, Kubo Y (2000) Novel fungal transcriptional activators, Cmr1p of *Colletotrichum lagenarium* and Pig1p of *Magnaporthe grisea*, contain Cys_2His_2 zinc finger and $Zn(II)_2Cys_6$ binuclear cluster DNA-binding motifs and regulate transcription of melanin biosynthesis genes in a developmentally specific manner. Mol Microbiol 38:940–954

Vienken K, Fischer R (2006) The $Zn(II)_2Cys_6$ putative transcription factor NosA controls fruiting body formation in *Aspergillus nidulans*. Mol Microbiol 61:544–554

Vienken K, Scherer M, Fischer R (2005) The $Zn(II)_2Cys_6$ putative *Aspergillus nidulans* transcription factor repressor of sexual development inhibits sexual development under low-carbon conditions and in submersed culture. Genetics 169:619–630

Wiemann P, Willmann A, Straeten M, Kleigrewe K, Beyer M, Humpf HU, Tudzynski B (2009) Biosynthesis of the red pigment bikaverin in *Fusarium fujikuroi*: genes, their function and regulation. Mol Microbiol 72:931–946

Wight WD, Kim KH, Lawrence CB, Walton JD (2009) Biosynthesis and role in virulence of the histone deacetylase inhibitor depudecin from *Alternaria brassicicola*. Mol Plant Microbe Interact 22:1258–1267

Young C, McMillan L, Telfer E, Scott B (2001) Molecular cloning and genetic analysis of an indole-diterpene gene cluster from *Penicillium paxilli*. Mol Microbiol 39:754–764

Yu J, Chang P-K, Ehrlich KC, Cary JW, Bhatnagar D, Cleveland TE, Payne GA, Linz JE, Woloshuk CP, Bennett JW (2004) Clustered pathway genes in aflatoxin biosynthesis. Appl Environ Microbiol 70:1253–1262

Yuan GF, Fu YH, Marzluf GA (1991) *nit-4*, a pathway-specific regulatory gene of *Neurospora crassa*, encodes a protein with a putative binuclear zinc DNA-binding domain. Mol Cell Biol 11:5735–5745

Yuan XL, Roubos JA, van den Hondel CA, Ram AF (2008) Identification of InuR, a new $Zn(II)_2Cys_6$ transcriptional activator involved in the regulation of inulinolytic genes in *Aspergillus niger*. Mol Genet Genomics 279:11–26

Zhang S, Monahan BJ, Tkacz JS, Scott B (2004) Indole-diterpene gene cluster from *Aspergillus flavus*. Appl Environ Microbiol 70:6875–6883